Not even wrong: Comment by Wagg et al.

Wagg, Cameron ; Barry, Kathryn E ; O'Brien, Michael J ; McKenzie-Gopsill, Andrew ; Roscher, Christiane ; Eisenhauer, Nico ; Schmid, Bernhard

DOI: https://doi.org/10.1002/ecy.2805

Not even wrong: Comment by Wagg et al.

Cameron Wagg[1], Kathryn E. Barry[2,3], Michael J. O'Brien[4], Andrew McKenzie-Gopsill[5], Christiane Roscher[3,6], Nico Eisenhauer[3,7], Bernhard Schmid[8]*

[1] Fredericton Research and Development Center, Agriculture and Agri-Food Canada, 850 Lincoln Road, Fredericton, New Brunswick, E3B 4Z7, Canada

[2] Systematic Botany and Functional Biodiversity, Institute of Biology, Leipzig University; Johannisallee 21, 04103 Leipzig, Germany

[3] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Deutscher Platz 5e, 04103 Leipzig, Germany.

[4] Área de Biodiversidad y Conservación, Universidad Rey Juan Carlos, c/Tulipán s/n, E-28933 Móstoles, Spain

[5] Charlottetown Research and Development Center, Agriculture and Agri-Food Canada, 440 University Ave., Charlottetown, Prince Edward Island, C1A 7Z5, Canada

[6] Department of Physiological Diversity, UFZ - Helmholtz Centre for Environmental Research, Permoserstr. 15, Leipzig 04318, Germany

[7] Leipzig University, Institute of Biology, Deutscher Platz 5e, 04103 Leipzig, Germany

[8] Department of Geography, University of Zürich, Winterthurerstrasse 190, 8057, Zürich, Switzerland

* Author for correspondence

Main text

A recent paper by Pillai and Gouhier (2019) (PG) in Ecology argues that biodiversity–ecosystem functioning (BEF) calculations of biodiversity effects using the additive partitioning approach introduced by Loreau and Hector (2001) (LH) are fundamentally flawed. PG claim that the null expectation of BEF research, and thus of the additive partitioning approach, stating that in the absence of biodiversity effects the performance of a mixture should equal the average performance of its component species' monocultures, is an extension of the neutral theory of species coexistence (which strictly speaking would require equal monoculture yields not required by LH). Further, they argue that a more reasonable null expectation based on coexistence theory (CT) would be *"for species in mixtures to coexist and by some form of niche partitioning overyield."* Based on these assertions, PG conclude that *"overyielding of ecosystem properties should be a natural outcome of coexistence"* and thus past BEF research has overestimated positive effects of biodiversity, for which this natural outcome could not be considered. Although BEF and CT share related ecological concepts, their different goals, origins, and intended uses prevent us from directly substituting their applications and implications. We provide a brief background on some fundamental concepts that underpin BEF research to elucidate potential misconceptions that may have led PG to their conclusion that, based on CT, the null expectation of BEF research used by LH is flawed. Despite the differences between BEF and CT research origins and approaches, their commonalities have been recognized early on (Harper 1977, Vandermeer 1981, Loreau 2004) and require further empirical investigation that may lead to novel opportunities to theoretically predict, and empirically measure, how coexistence mechanisms may contribute to observed effects of biodiversity on ecosystem functioning in nature.

PG claim that findings of enhanced ecosystem functioning in more biodiverse systems based on the additive partitioning approach introduced by LH are inflated due to the

assumption of neutrality between species. They claim that the null expectation of no effect on ecosystem functioning with increasing species richness, while holding community density constant, is flawed because as richness increases these species must coexist by occupying different niches. Here we argue that this assertion may not apply for three reasons:

1. The history of BEF research and null expectations of no effect of additional species on ecosystem function has evolved from agricultural science and plant population biology independently of CT.
2. Although long-term stable coexistence may often be associated with overyielding, there can be coexistence without overyielding.
3. There can be overyielding without stable coexistence (Turnbull et al. 2013), and it is possible to create short-term communities of higher richness and productivity than possible in the long term (Pfisterer et al. 2004).

1. BEF research evolved independently of CT and thus is subject to different null expectations

BEF research started with the simple question about potential effects of species loss from an ecosystem on its functioning, often measured as biomass yield. For this, plant species richness was experimentally manipulated by holding total sowing or planting density constant to avoid confounding species richness and community density (substitutive as opposed to additive design; Schmid et al. 2002a). These experiments commonly found that plant species richness–yield relationships were positive but with declining increases as more and more species were assembled (e.g. Tilman et al. 1996, Hector et al. 1999; meta-analysis in Balvanera et al. 2006). This generated some debate at the time as to whether these positive effects of plant diversity on productivity were simply due to the increasing probability of including one, or a few, highly productive species in more diverse systems (e.g., Huston

74 1997, Wardle 2016, Eisenhauer et al. 2016). Note that this requires the so-called rule of
75 constant final yield (the first law of plant population biology according to Harper 1977),
76 which allows these highly productive species to approach the same biomass in mixtures as in
77 monoculture, even though they are sown at lower density in mixture. The initial skepticism
78 against the possible existence of positive BEF relationships was in part related to theory
79 which predicted negative relationships between ecosystem complexity and stability (May
80 1972). A further reason of skepticism were the expectations of early agricultural botanists
81 that, at least under ideal conditions, the best-performing monoculture should always be more
82 productive than the best mixture (no transgressive overyielding of mixtures, Harper 1977).
83 Assuming that BEF relationships must be positive by definition due to CT (as understood by
84 PG) is thus counter to these historical expectations.
85 This research was based, to a large extent, on the pioneering work of de Wit (1960, de
86 Wit et al 1966). De Wit introduced the replacement series experimental design, where total
87 density is kept constant but the mixing ratio of two species is varied. Monocultures of species
88 are used as a reference for identifying the optimal sward-mixing ratio along a replacement
89 series. This design is based on the null expectation that if there are no effects of mixing,
90 individuals of different species are competitively equivalent and thus mixtures should yield
91 as predicted by adding the monocultures in the proportions given by the frequency of the
92 species in mixture. Thus, classic replacement series assess the effect of increasing the sown
93 density of a species '*i*', while equally reducing the sown density of a species '*j*' to maintain a
94 constant sowing density (Fig. 1a). If the null expectation is that the yield per individual of a
95 species is constant in constant community density, then a species contribution to yield is its
96 monoculture yield weighted by its sown density proportion (Fig. 1b). The sum of the species
97 relative yields is the relative yield total (*RYT*), and if species are in mixture with a constant
98 community density, then the null hypothesis is that the $RYT = 1$. According to this first

scenario in Fig. 1, if all species were identical and density constant between mixtures and monocultures, then increasing species richness should have no effect on yield and no overyielding should occur ($RYT = 1$ for all species richness levels).

Of course, $RYT = 1$ is almost never the observed result in field trials as various ecological mechanisms can be at play. Some of these ecological mechanisms may be related to CT and the limiting-similarity and competitive-exclusion hypotheses (MacArthur & Levins 1967, Chesson 2000, Weiner and Freckleton 2010, Barabás et al. 2018). For instance, by reducing the density of species *i* in mixture, it may become released from intraspecific competition that could lead to increased yield per individual of species *i*. If the other species *j* does not suffer more from *i* than *i* gains from *j*, or if even both species behave in the described manner in mixture, i.e. if overall intraspecific competition is stronger than interspecific competition, then $RYT > 1$ (Fig. 1: scenarios two and three). As long as there is no facilitation between the two species, a maximum $RYT = 2$ can be achieved in two-species mixtures according to the rule of constant final yield (Fig. 1: scenario two; Schmid et al. 2002b). If one or both species facilitate(s) the growth of the other—such as in the case of legumes facilitating neighboring species (Spehn et al. 2005, Temperton et al. 2007), and if there would be no interspecific competition at all, then $RYT > 2$ would theoretically be possible.

Whatever the species richness, according to scenario two the expected mixture yield is the sum of all the component species' monoculture yields. Such a situation with no niche overlap and no interspecific competition between species is of course highly unrealistic but important to consider in the context of PG's new null expectation for biodiversity effects, because it reflects this null expectation in its most extreme form. They argue that the null expectation of BEF research should be that mixtures are overyielding to the degree to which species niches do not overlap and thus do not compete for the same resources. In this case,

niche overlap between species under the null expectation could be measured via *RYT* and, by doing this in all possible two-species mixtures, null expectations for mixtures of more than two species could be derived. However, to tease apart the two-way interactions from the higher-order interactions would require new experiments that include replicate monocultures, replicate 2-species mixtures with all pairs of species and replicated mixtures with higher levels of diversity, a heroic effort that has not been done thus far. Not surprisingly, the real biodiversity effect of interest calculated by PG is generally negative, because they use no correction for multiple niche overlap in mixtures of three or more species. Their null expectation is thus fundamentally different from the common null expectation used in LH, which corresponds to the first scenario in Fig. 1. Here the expected mixture yield is not the sum but the average of all the component species' monoculture yields and thus increasing species richness does not change yield.

The additive partitioning of biodiversity effects introduced by LH allowed for the first time to quantify different ways in which relative yields of species could combine to total yield, something that previously was done in more qualitative ways e.g. by the replacement-series approach of de Wit (1960). Additive partitioning works with relative yields to separate the net biodiversity effect (difference between mixture yield and the average of monoculture yields) into two additive components called the 'complementarity effect' (CE) and the 'selection effect' (SE). The additive partitioning was developed to assess early BEF debates as to whether positive net effects (NE=CE+SE) of biodiversity were driven more or less by particularly productive species (SE) or by a generally increased species performance in mixture (CE). It was not intended to tease apart the density-dependent competitive effects, or fitness differences, between species pairs used in CT, or to provide direct evidence of any one particular ecological mechanism underpinning BEF relationships as PG incorrectly assume. It was already known prior to the development of the additive partitioning that the use of

149 relative yields confounds intraspecific and interspecific density effects (see Connolly 1986,
150 Jolliffe 2000).

151 Application of the LH additive partitioning scheme increases the utility of
152 replacement series results. Originally, the aim of the *RYT* was to compare species of similar
153 monoculture performance in a replacement series. This similarity in monoculture yield
154 simplifies the interpretation of $RYT > 1$, because if species monocultures are similar in yield,
155 a $RYT > 1$ will always require a contribution of both species. In this case, the overyielding of
156 the mixture is mostly or fully due to the CE from the additive partitioning method. However,
157 if a more productive species is mixed with a less productive species, $RYT > 1$ can but does
158 not have to be achieved with the SE alone. It is here where the usefulness of the additive
159 partitioning approach can be seen: it distinguishes this case by assigning the NE of $RYT > 1$
160 to SE, where the mixture becomes essentially a monoculture of the more productive species,
161 from other cases where both species contribute to the biodiversity effect. In fact, if two
162 species differ in monoculture yields, there is a large range of possible contributions of SE and
163 CE that lead to NE > 0 and $RYT > 1$ (see Fig. 1). Similar arguments apply to cases with RYT
164 < 1 (Loreau and Hector 2001).

165 Considering overyielding in species mixtures originates from assessing the optimal
166 sward mixing in managed grasslands and whether it is better to plant monocultures or
167 mixtures (e.g. de Wit 1960). The use of the additive partitioning of biodiversity effects to
168 tease apart the relative contributions of species in mixtures as well as to explain why mixtures
169 perform differently than the average monoculture is useful for the analysis of BEF
170 experiments. The arguments of PG overlooked this rich history and the intended use for
171 assessing yield that does not need to be directly connected to the long-term coexistence of
172 species. A further very useful aspect of the additive partitioning method is that it applies not
173 only to two-species mixtures but also to any other species richness levels. The only

174 requirement is that species yields can be separated in mixtures and that all species can also be
175 grown in monoculture.
176
177 **2. There can be coexistence without overyielding**
178 The fact that CEs and SEs do not have a direct mechanistic explanation in terms of
179 coexistence or niche theory may be one of the frustrations of PG. Indeed, a common
180 misconception is that complementarity effects should be directly equivalent to resource
181 complementarity or spatial complementarity (Barry et al. 2019). While this is an interesting
182 hypothesis, it has received little empirical support so far. For example, in a forest experiment,
183 physical crown complementarity was actually more closely related to SEs than to CEs
184 (Williams et al. 2017), perhaps reflecting a competitive trait hierarchy (Kunstler et al. 2012).
185 Disconnects between species coexistence and the mechanisms by which species contribute to
186 overyielding reveal the difficulty in logically substituting the concepts of one into the other
187 and may be based on misconceptions of the origins and objectives of BEF research for the
188 following reasons.
189 First, although positive effects of biodiversity are commonly observed in both
190 experimental and natural contexts and across different ecosystems (Balvanera et al. 2006,
191 Duffy et al. 2017), the experimental mixing of species does not always result in overyielding.
192 In rare cases, antagonistic interactions among species have also been observed in biodiversity
193 experiments leading to mixtures having $RYT < 1$. This occurs if overall inter-specific
194 competition is stronger than intraspecific competition in a mixture (Fig. 1: scenario four). For
195 example, this has been observed in bacterial communities, where species increase the
196 production of allelopathic toxins in mixtures, thus leading to a strongly negative biodiversity
197 effect without the competitive exclusion of species during the course of an experiment
198 (Jousset et al. 2011, Becker et al. 2012). Other situations resulting in $RYT < 1$ can result from

199 positive density-dependence, for example the lack of attracting mutualists by a species at
200 lower densities—such as in the case where the performance of a species is dependent on
201 positive plant–soil feedback effects (van der Putten et al. 2013). While in these cases both
202 species may suffer in a two-species mixture, $RYT < 1$ can also result in situations where an
203 unproductive species *j* suppresses a more productive species *i*, even though *j* may be less
204 suppressed by *i* than by itself (Loreau & Hector 2001).

205 Second, if similar species are expected to compete more strongly for common
206 resources, then it may also be expected that due to their similarity the species also have
207 relatively equal competitive effects on each other (Ebeling et al. 2014). Thus, in mixtures of
208 highly similar species, neither will have a large enough competitive advantage to exclude the
209 other due to their similar fitness (Aarssen 1983, Hubbell 2001, Adler et al. 2007, Mayfield
210 and Levine 2010). The consequence of species being competitively equivalent with a high
211 level of niche overlap could again result in species coexisting without overyielding. In such
212 cases, it is clear that statements such as *"overyielding of ecosystem properties should be a*
213 *natural outcome of coexistence"* by PG have overlooked the fact that species in mixture can
214 interact and coexist in a way that does not result in overyielding.

215

216 **3. There can be overyielding without stable coexistence**

217 The hypothesized mechanisms that underpin the empirical observations of positive
218 BEF relationships and the hypothesized drivers of coexistence run parallel to each other
219 (Loreau 2004, Loreau 2010, Caroll et al. 2011, Turnbull et al. 2013, Turnbull et al. 2016).
220 Such parallels may give the impression that the approaches are aimed at achieving the same
221 objective and may be easily intermixed as done by PG. However, CT and BEF relationships
222 are not completely comparable, and long-term stable coexistence is not a necessary condition
223 for enhanced ecosystem functioning in more diverse systems as implied by PG.

First, CT and the additive partitioning method are used at two different temporal scales. Additive partitioning is applied to a state variable usually within a single growing season. In contrast, classic coexistence models are based on per capita growth rates modeled over many demographic turnover events in a community at stable equilibrium and are therefore process variables. Commonalities between a complementarity effect in BEF research and competitive interactions from CT may be more easily drawn in biodiversity experiments involving trees, where the yearly growth of individuals can be followed (Huang et al. 2018), or in experiments involving perennial herbaceous plants if we replace 'per capita growth rate' with 'species biomass accumulation rate' within a single growing season, and if we also assume that the species monoculture yield is an indication of species carrying capacity *K*. The use of monoculture yields to infer *K* is not unreasonable because according to the law of constant final yield (monocultures sown at a very high density undergo 'self-thinning' until the constant final yield is met, Harper 1977, Weiner and Freckleton 2010). This constant final yield can also be reached with much lower density, where individual plants grow to be larger (Harper 1977). The concept of constant final yield in monocultures may be considered akin to the concept of a species carrying capacity (*K*) used in species competition-coexistence frameworks (Westoby 1981). Thus, plants may 'adjust' their 'per capita growth rate' within a single cohort to compensate for density variations to achieve the constant final yield (*K*). This is an observed phenomenon in various plant species populations, where individual plants at lower intraspecific densities grow larger, likely due to greater resource availability per individual (Bazzaz and Harper 1976, Malmberg and Smith 1982, Roscher and Schumacher 2016). However, effects of species richness on thinning and constant final yield in mixtures have barely been explored and are complicated (He et al. 2005, Roscher et al. 2007), further demonstrating the complexity of relating BEF research to CT.

CEs and SEs often change over many years in long-term biodiversity experiments (Cardinale et al. 2007, Marquard et al. 2009, Reich et al. 2012, Huang et al. 2018). Temporal changes in CEs (and *RYT*) can be generated from coexistence models (Turnbull et al. 2013). However, these theoretical models demonstrate that enhanced ecosystem functioning can occur even when long-term stable coexistence is not possible. Turnbull et al. (2013) found that communities can overyield even when stabilizing niche differences (*sensu* Chesson 2000) cannot overcome fitness differences and therefore allow for stable coexistence. This 'transient' overyielding may occur frequently in natural communities, where environmental conditions change and may lead to long-term unstable persistence rather than long-term stable coexistence (Wagg et al. 2017). Furthermore, nutrient addition sometimes can increase overyielding and simultaneously decrease the number of species that are able to coexist, indicating that enhanced ecosystem functioning with increased species richness and coexistence in these systems may be driven by different mechanisms (Fox and Harpole 2008, Harpole et al. 2016). Therefore, the application of the coexistence models to diverse perennial grasslands, where much of the BEF research originated, including the use of the additive partition method, is extremely challenging as species vary in their demographic turnover rates. There is no constant per capita growth rate considering that plant yield–density relationships follow the low of constant final yield, and that plant communities can be highly diverse (more than species pairs), resulting in complex higher-order interactions that are difficult to parameterize.

Conclusions

Coexistence mechanisms based on niche partitioning and fitness differences are important ecological mechanisms that can relate to positive effects of biodiversity and are useful for deriving empirically testable hypotheses for future studies on the mechanisms

underlying BEF relationships. However, the proposed redefining of the null hypothesis by PG, claiming that coexistence always implies "trivial" positive biodiversity effects against which "true" biodiversity effects would have to be evaluated is circular in its own way as indicated by the use of the trivial component of biodiversity effects as a measure of coexistence. In other words, this null hypothesis defines a biodiversity effect as proof of no effect. Furthermore, the use of classic coexistence models for predicting species competitive outcomes at higher levels of diversity (with more than two species) is limited due to higher-order interactions that become more complex across increasing levels of species richness (Levine et al. 2017, Barabás et al. 2018). The additive partitioning of biodiversity effects is also not theoretically limited by first having to parameterize all species' pairwise and higher order interactions in diverse plant mixtures. While the use of CT indeed provides many insights for understanding ecological mechanisms that support biodiversity in nature, its use in practical and applied settings, such as in agricultural ecosystems, where species mixtures are sown at predefined densities and responses are quantified within a single growing season, may not be as useful as relative yields and the additive partitioning method.

References

Aarssen, L.W. 1983. Ecological combining ability and competitive combing ability in plants: toward a general evolutionary theory of coexistence in systems of competition. The American Naturalist 122:707-731.

Adler, P., J. HilleRisLambers and J.M. Levine. 2007. A niche for neutrality. Ecology Letters 10:95-104.

Balvanera, P., A.B. Pfisterer, N. Buchmann, J.S. He, T. Nakashizuka, D. Raffaelli and B. Schmid. 2006. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. Ecology Letters 9:1146-1156.

299 Barabás, G., R.D. Andrea and S.M. Stump. 2018. Chesson's coexistence theory. Ecological
300 Monographs 88:277-303.
301 Barry, K.E., et al. 2019. The future of complementarity: disentangling causes from
302 consequences. Trends in Ecology and Evolution 34:167-180.
303 Bazzaz, F.A. and J.L. Harper. 1976. Relationship between plant weight and numberss in
304 mixed populations of *Sinapsis alba* (L.) Rabenh. and *Lepidium sativum* L. Journal of
305 Applied Ecology 13:211-216.
306 Becker, J., N. Eisenhauer, S. Scheu and A. Jousset. 2012. Increasing antagonistic interactions
307 cause bacterial communities to collapse at high diversity. Ecology Letters 15:486-474.
308 Cardinale, B. J., J. P. Wright, M. C. Cadotte, I. T. Carroll, A. Hector, D. S. Srivastava, M.
309 Loreau and J. J. Weiss. 2007. Impacts of plant diversity on biomass increase through time
310 because of species complementarity. Proceedings of the National Academy of Sciences,
311 USA 104:18123-18128.
312 Carroll, I.T., B.J. Cardinale and R.M. Nisbet. 2011. Niche and fitness differences relate the
313 maintenance of diversity to ecosystem functioning. Ecology 92:1157-1165.
314 Chesson, P. 2000. Mechanisms of maintenance of species diversity. Annual Review of
315 Ecology, Evolution, and Systematics, 31, 343–366.
316 Connolly, J. 1986. On difficulties with the replacement-series methodology in mixture
317 experiments. Journal of Applied Ecology 23:125-137.
318 de Wit, C. T. 1960. On competition. Verslag Landbouwkundige Onderzoek 66.8:1-82.
319 de Wit, C. T., P. G. Tow, and G. C. Ennik. 1966. Copetition between legumes and grasses.
320 Verslag Landbouwkundige Onderzoek 687:1-30.
321 Harpole, W.S. and J.W. Fox. 2008. Revealing how species loss affects ecosystem function:
322 the trait-based Price equation partition. Ecology 89:269-279.

323 Duffy, J.E., C.M. Goodin and B.J. Cardinale. 2017. Biodiversity effects in the wild are
324 common and as strong as key drivers of productivity. Nature 549:261-264.
325 Ebeling, A. Et al. 2014. A trait-based experimental approach to understand the mechanisms
326 underlying biodiversity–ecosystem functioning relationships. Basic and Applied
327 Ecology 15: 229-240.
328 Eisenhauer, N., D. B. Barnes, S. Cesarz, D. Craven, O. Ferlian, F. Gottschall, J. Hines, A.
329 Sendek, J. Siebert, M. P. Thakur and M. Türke. 2016. Biodiversity-ecosystem function
330 experiments reveal the mechanisms underlying the consequences of biodiversity
331 change in real world ecosystems. Journal of Vegetation Science 27:1061-1070.
332 Harper, J.L. 1977. Population Biology of Plants. Academic Press, London.
333 Harpole, W.S., et al. 2016. Addition of multiple limiting resources reduces grassland
334 diversity. Nature 537:93-96.
335 He, J.-S., K.S. Wolfe-Bellin, B. Schmid and F.A. Bazzaz. 2005. Density may alter diversity-
336 productivity relationships in experimental plant communities. Basic and Applied
337 Ecology 6:505–517.
338 Hector, A., et al. 1999. Plant diversity and productivity experiments in European grasslands.
339 Science 286:1123–1127.
340 Huang, Y., et al. 2018. Impacts of species richness on productivity in a large-scale
341 subtropical forest experiment. Science 362:80-83.
342 Hubbell, S.P. 2001. The unified neutral theory of biodiversity and biogeography (MBP-32).
343 Princeton University Press.
344 Huston, M.A. 1997. Hidden treatments in ecological experiments: re-evaluating the
345 ecosystem function of biodiversity. Oecologia 110:449-460.
346 Jolliffe, P. A. 2000. The replacement series. Journal of Ecology 88:371-385.

347 Jousset, A., B. Schmid and N. Eisenhauer. 2011. Genotypic richness and dissimilarity
348 opposingly affect ecosystem functioning. Ecology Letters 14:537-545.
349 Kunstler, G., S. Lavergne, B. Courbaud, W. Thuiller, G. Vieilledent, N.E. Zimmermann, J.
350 Kattge, D.A. Coomes. 2012. Competitive interactions between forest trees are driven
351 by species' trait hierarchy, not phylogenetic or functional similarity: Implications for
352 forest community assembly. Ecology Letters 15:831–840.
353 Levine, J.M., J. Bascompte, P.B. Adler and S. Allesina. 2017. Beyond pairwise mechanisms
354 of species coexistence in complex communities. Nature 546:56-64.
355 Loreau, M. 2004. Does functional redundancy exist? Oikos 104:606-611.
356 Loreau, M. 2010. Linking biodiversity and ecosystems: towards a unifying ecological theory.
357 Philosophical Transactions of the Royal Scoiety B 365:49-60.
358 Loreau, M., and A. Hector. 2001. Partitioning selection and complementarity in biodiversity
359 experiments. Nature 412:72 76.
360 MacArthur, R. H., and R. Levins. 1967. The limiting similarity, convergence, and divergence
361 of coexisting species. American Naturalist 101:377–385.
362 Malmberg, C. and H. Smith. 1982. Relationship between plant weight and density in mixed
363 populations of *Medicago sativa* and *Trifolium pretense*. Oikos 38:365-368.
364 Marquard, E., A. Weigelt, V. M. Temperton, C. Roscher, J. Schumacher, N. Buchmann, M.
365 Fischer, W. W. Weisser and B. Schmid. 2009. Plant species richness and functional
366 composition drive overyielding in a six-year grassland experiment. Ecology 90:3290-
367 3302.
368 May, R.M. 1972. Will a large complex system be stable? Nature 238:413-414.
369 Mayfield, M. and J. Levine. 2010. Opposing effects of competitive exclusion on the
370 phylogenetic structure of communities. Ecology Letters 13:1085-1093.

371 Pfisterer , A.B., J. Joshi, B. Schmid and M. Fischer. 2004. Rapid decay of diversity-
372 productivity relationships after invasion of experimental plant communities. Basic and
373 Applied Ecology 5:5-14.
374 Pillai, P. and T.C. Gouhier. 2019. Not even wrong: the spurious measurement of
375 biodiversity's effects on ecosystem functioning. Ecology (in press)
376 https://doi.org/10.1002/ecy.2645
377 Reich, P.B., D. Tilman, F. Isabell, K. Mueller, S. E. Hobbie, D. F. B. Flynn and N.
378 Eisenhauer. 2012. Impacts of biodiversity loss escalate through time as redundancy fades.
379 Science 336:589-592.
380 Roscher, C. and J. Schumacher. 2016. Positive diversity effects on productivity in mixtures
381 of arable weed species as related to density-size relationships. Journal of Plant Ecology
382 9:792-804.
383 Roscher, C., J. Schumacher, W.W. Weisser, B. Schmid and E.D. Schulze. 2007. Detecting
384 the role of individual species for overyielding in experimental grassland communities
385 composed of potentially dominant species. Oecologia. 154:535–549.
386 Schmid, B., A. Hector, M.A. Huston, P. Inchausti, I. Nijs, P.W. Leadley and D. Tilman.
387 2002a. The design and analysis of biodiversity experiments. Pages 61-75 *in* M. Loreau,
388 S. Naeem and P. Inchausti, editors. Biodiversity and ecosystem functioning: synthesis
389 and perspectives. Oxford University Press, Oxford, UK.
390 Schmid, B., J. Joshi and F. Schläpfer. 2002b. Emperical evidence for biodiversity-ecosystem
391 functioning relationships. Pages 120–150 *in* A. Kinzig, S. Pacala and D. Tilman,
392 editors. The Functional consequences of biodiversity: empirical progress and
393 theoretical extensions (MPB-33). Princeton University Press, New Jersey, USA.
394 Spehn, E.M., et al. 2005. Ecosystem effects of biodiversity manipulations in European
395 grasslands. Ecological Monographs 75:37-63.

396 Temperton, V. M., P. N. Mwangi, M. Scherer-Lorenzen, B. Schmid, and N. Buchmann.
397 2007. Positive interactions between nitrogen-fixing legumes and four different
398 neighbouring species in a biodiversity experiment. Oecologia 151:190-205.
399 Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by
400 biodiversity in grassland ecosystems. Nature 379:718-720.
401 Turnbull, L.A., J. M. Levine, M. Loreau and A. Hector. 2013. Coexistance, niches and
402 biodiversity effects on ecosystem functioning. Ecology Letters 16:116-127.
403 Turnbull, L.A., F. Isbell, D.W. Purves, M. Loreau and A. Hector. 2016. Understanding the
404 value of plant diversity and ecosystem functioning through niche theory. Proceedings
405 of the Royal Society B 283:20160536.
406 Vandermeer, J. 1981. The interference production principle: an ecological theory for
407 agriculture. BioScience 31:361-364.
408 van der Putten, W.H., et al. 2013. Plant-soil feedbacks: the past, the present and future
409 challenges. Journal of Ecology 101:265–276.
410 Wagg, C., M.J. O'Brien, A. Vogel, M. Scherer-Lorenzen, N. Eisenhauer, B. Schmid and A.
411 Weigelt. 2017. Plant diversity maintains long-term ecosystem productivity under
412 frequent drought by increasing short-term variation. Ecology 98:2952-2961.
413 Wardle, D.A. 1999. Is "sampling effect" a problem for experiments investigating
414 biodiversity-ecosystem function relationships? Oikos 87:403-407.
415 Wardle, D.A. 2016. Do experiments exploring plant diversity–ecosystem functioning
416 relationships inform how biodiversity loss impacts natural ecosystems? 27:646-653.
417 Weiner, J. and R.P. Freckleton. 2010. Constant final yield. Annual Review of Ecology,
418 Evolution, and Systematics 41:173-192.
419 Westoby, M. 1981. The place of the self-thinning rule in population dynamics. The American
420 Naturalist 118:581-587.

421 Williams, L.J., A. Paquette, J. Cavender-Bares, C. Messier and P.B. Reich. 2017. Spatial
422 complementarity in tree crowns explains overyielding in species mixtures. Nature
423 Ecology and Evolution 1:0063.
424

Figure 1

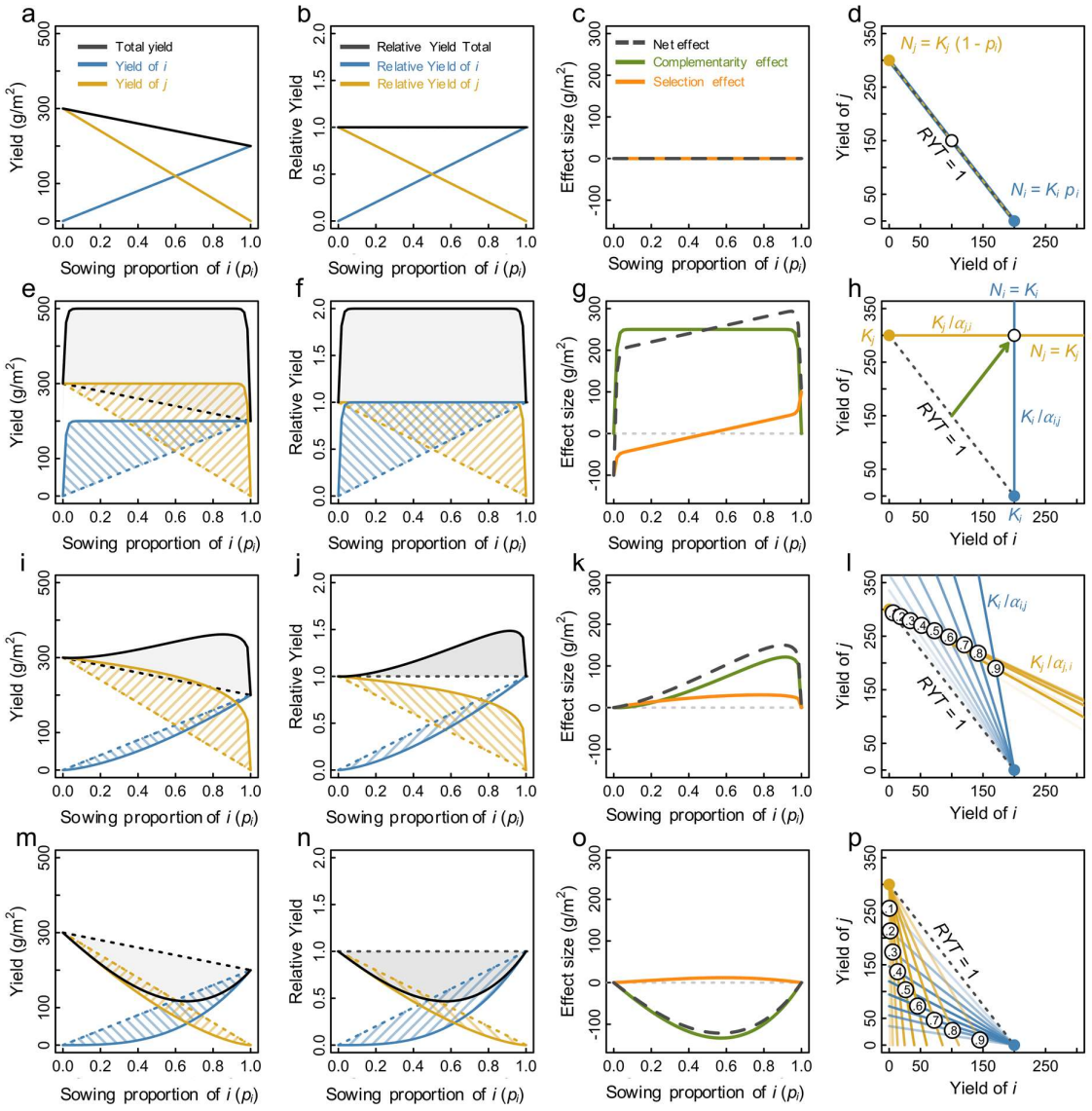

Figure 1. Diagram illustrating the classic 'de Wit' style of mixing two species ('*i*' and '*j*' with monoculture yields of 300 and 200 g/m^2 respectively) along a replacement series design with four different hypothetical scenarios (figure rows). The first scenario (a–d) is the null expectation for biodiversity effects where there are intra- and interspecific interactions are equal and the yield per sown density corresponds to simple mixing of the monoculture yields. The second scenario (e–h) illustrates when there are only intraspecific and no interspecific interactions, thus species achieve their constant final yield at all except the very lowest

sowing densities, which is an extreme form of a biodiversity effect and the most extreme case of the null expectation of Pillai and Gouhier (2019). The third scenario (i–l) illustrates where the less productive species *i* (blue) yields less than expected in mixtures, but is overcompensated for by the more productive species *j* (light-brown) yielding more than expected in mixtures leading to a positive selection and complementarity effect at higher sowing proportions of species *i* to *j*. The fourth scenario (m–p) illustrates the case when both species yield less than expected in mixture resulting in a negative complementarity effect (here the less productive species *i* yields proportionally slightly less than the more productive species *j*, thus, producing a slightly positive selection effect). For each scenario we have illustrated the results in the figure columns of yield (a, e, i, m), relative yield (b, f, j, n), net biodiversity effects (NE, black-dashed line), complementarity effects (green) and selection effects (orange; c, g, k, o) and a projection of the outcomes onto species isoclines assuming stable coexistence is achieved at each sowing combination (d, h, l, p, *sensu* Loreau 2004). For yield and relative-yield panels the dotted lines indicate the expected yields and relative yields under the null hypothesis that the yield of an individual is density independent for both species (first scenario). The shaded region in the yield and relative-yield figures indicates the deviation away from the null expectation. For panels showing NEs, CEs and SEs the lightly dotted grey line highlights NE = 0 (no effect). In (d) under the null hypotheses species yield at equilibrium is linearly proportional to its relative sown density. Here the open point simply highlights a 1:1 sowing ratio, but the two species can stably coexist anywhere along the line. In (h) the change in the proportion sown of one species has no effect (no slope) on the yield of the other and species will always achieve an equilibrium yield consistent with their monocultures regardless of their sown proportion. Here the dotted line indicates the expected relative yield total in the absence of any overyielding ($RYT = 1$) and the green arrow highlights the overyielding effect of the species in mixture. The open point indicates the

yields at which species coexist (i.e. at each sown proportion). In panel (l) species isoclines are shown for each sowing proportion of i from 0.1 to 0.9 in sequential steps of 0.1 (indicated with the open points highlighting the yields at which species coexist, i.e. at each sown proportion). More darkly colored lines indicate a stronger complementarity effect (distance away from $RYT = 1$). The same is shown in panel (p) as in panel (l) for the situation where species mixing results in under-yielding ($RYT < 1$).